INFORMATION PROCESSING & MANAGEMENT

An International Journal

Editor-in-Chief: **Tefko Saracevic**
Email: tefko@scils.rutgers.edu
Home page: http://www.scils.rutgers.edu/~tefko
Rutgers University

Volume 39 (2003) Contents and Author Index

doi:10.1016/S0306-4573(03)00060-8

Pergamon

INFORMATION PROCESSING & MANAGEMENT

– An International Journal –

(Incorporating *INFORMATION TECHNOLOGY*)

Email: ipm@scils.rutgers.edu

Website:http://www.elsevier.com/locate/infoproman

Editor-in-Chief

Tefko Saracevic
School of Communication, Information and Library Studies
Rutgers University
4 Huntington Street, New Brunswick, NJ 08903, U.S.A.
tefko@scils.rutgers.edu
http://www.scils.rutgers.edu/~tefko

Associate Editors

Nicholas J. Belkin
School of Communication, Information and Library Studies
Rutgers University
4 Huntington Street
New Brunswick, NJ 08901-1071, U.S.A.
nick@belkin.rutgers.edu

W. Bruce Croft
Department of Computer and Information Science
University of Massachusetts
Amherst, MA 01003, U.S.A.
croft@cs.umasss.edu

Associate Editor (Europe)

C. J. van Rijsbergen
Department of Computing Science
Lilybank Gardens
The University
Glasgow G12 8QQ, Scotland
keith@des.gla.ac.uk

Associate Editor (Book Reviews)

Amanda Spink
Associate Professor
School of Information Sciences & Technology
The Pennsylvania State University
120 S. Burrowes Street, 511 Rider Building
University Park, PA 16801-3857, U.S.A.
spink@ist.psu.edu

Founding Editor

Harold Borko
Graduate School of Education and Information Studies
University of California
102 South Hall
Los Angeles, CA 90024-1520, U.S.A.
hborko@ucla.edu

Editorial Board

Maristella Agosti
Padova University, Padova, Italy
agosti@dei.unipd.it

Tatiana Aparac
University J.J. Strossmayer
Osijek, Croatia
taparac@mudrac.ffzg.hr

David C. Blair
University of Michigan
Ann Arbor, MI, U.S.A.
dcblair@sils.umich.edu

Abraham Bookstein
University of Chicago
Chicago, IL, U.S.A.
a-bookstein@uchicago.edu

Pia Borlund
The Royal School of Library and Information Science
Aalborg, Denmark
pbj@db.dk

Michael Buckland
University of California
Berkeley, CA, U.S.A.
buckland@sims.berkeley.edu

Yves Chiaramella
University of Grenoble
Grenoble Cedex, France
chiara@imag.fr

Susan Dumais
Microsoft Research
Redmond, WA, U.S.A.
sdumais@microsoft.com

Raya Fidel
University of Washington
Seattle, WA, U.S.A.
fidelr@u.washington.edu

Edward A. Fox
Virginia Polytechnic Institute & State University
Blacksburg, VA, U.S.A.
fox@fox.cs.vt.edu

Jonathan Furner
University of California
Los Angeles, CA, U.S.A.
jfurner@ucla.edu

Donna Harman
National Institute of Standards & Technology
Gaithersburg, MD, U.S.A.
donna.harman@nist.gov

David Harper
The Robert Gordon University
Aberdeen, Scotland
djh@scms.rgu.ac.uk

William Hersh
Oregon Health Sciences University
Portland, OR, U.S.A.
hersh@ohsu.edu

Peter Ingwersen
The Royal School of Library and Information Science
Copenhagen, Denmark
pi@db.dk

Tetsuya Ishikawa
University of Library and Information Science
Taukuba, Japan
ishikawa@slis.tsukuba.ac.jp

Kalervo Järvelin
University of Tampere, Tampere, Finland
kalervo.jarvelin@uta.fi

Noriko Kando
National Institute of Informatics
Tokyo, Japan
kando@nii.ac.jp

Paul B. Kantor
Rutgers University
New Brunswick, NJ, U.S.A.
kantor@scils.rutgers.edu

Gehard Knorz
Fachhochschule Darmstadt
Darmstadt, Germany
knorz@www.iud.fh-darmstadt.de

Matthew Koll
6912 Nevis Road
Bethesda, MD, U.S.A.
mkoll@aol.com

Mun-Kew Leong
BIGontheNet, Singapore
MunKew.Leong@BIGontheNet.com

Elisabeth D. Liddy
Syracuse University, Syracuse, NY, U.S.A.
liddy@mailbox.syr.edu

John Lindsay
Kingston University
Kingston upon Thames, U.K.
lindsay@king.ac.uk

Gary Marchionini
University of North Carolina
Chapel Hill, NC, U.S.A.
marcus@ils.unc.edu

Richard S. Marcus
Massachusetts Institute of Technology
Cambridge, MA, U.S.A.
marcus@lids.mit.edu

Michel J. Menou
Consultant in Information Management
Les Rosiers sur Loire, France
mmenou@imaginet.fr

Sung Hung Myaeng
Chungnam National University
Taejon, Korea
shmyaeng@cs.chungnam.ac.kr

Fausto Rabitti
Consiglio Nazionale delle Ricerche
Pisa, Italy
F. Rabitti@cnuce.cnr.it

Edie M. Rasmussen
University of Pittsburgh
Pittsburgh, PA, U.S.A.
erasmus@sis.pitt.edu

Stephen E. Robertson
Microsoft Research Ltd., Cambridge, U.K.
ser@microsoft.com

Mark Sanderson
University of Sheffield
Sheffield, South Yorkshire, U.K.
m.sanderson@sheffield.ac.uk

Fabrizio Sebastiani
Istituto di Scienza e Tecnologie dell'Informazione,
Consiglio Nazionale delle Ricerche Pisa, Italy
Fabrizio@iei.pi.cnr.it

Alan F. Smeaton
Dublin City University
Dublin, Ireland
Alan.Smeaton@dcu.ie

Karen Sparck Jones
University of Cambridge
Cambridge, U.K.
sparckjones@cl.cam.ac.uk

Pertti Vakkari
University of Tampere
Tampere, Finland
Pertti.Vakkari@uta.fi

Author enquiries: For enquiries relating to the submission of articles (including electronic submission where available) please visit Elsevier's Author Gateway at http://authors.elsevier.com. The Author Gateway also provides the facility to track accepted articles and set up e-mail alerts to inform you of when an article's status has changed, as well as detailed artwork guidelines, copyright information, frequently asked questions and more.
Contact details for questions arising after acceptance of an article, especially those relating to proofs, are provided after registration of an article for publication.

Publication information: *Information Processing & Management* (ISSN 0306-4573). For 2003, volume 39 is scheduled for publication. Subscription prices are available upon request from the Publisher or from the Regional Sales Office nearest you or from this journal's website: (http://www.elsevier.com/locate/infoproman). Further information is available on this journal and other Elsevier products through Elsevier's website: (http://www.elsevier.com). Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by standard mail (surface within Europe, air delivery outside Europe). Priority rates are available upon request. Claims for missing issues should be made within six months of the date of dispatch.

Orders, claims, and journal enquiries: please contact the Customer Service Department at the Regional Sales Office nearest you:
Orlando: Elsevier, Customer Service Department, 6277 Sea Harbor Drive, Orlando, FL 32887-4800, USA; phone: (+1) (877) 8397126 [toll free number for US customers], or (+1) (407) 3454020 [customers outside US]; fax: (+1) (407) 3631354; e-mail: usjcs@elsevier.com
Amsterdam: Elsevier, Customer Service Department, PO Box 211, 1000 AE Amsterdam, The Netherlands; phone: (+31) (20) 4853757; fax: (+31) (20) 4853432; e-mail: nlinfo-f@elsevier.com
Tokyo: Elsevier, Customer Service Department, 4F Higashi-Azabu, 1-Chome Bldg, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106-0044, Japan; phone: (+81) (3) 5561 5037; fax: (+81) (3) 5561 5047; e-mail: jp.info@elsevier.com
Singapore: Elsevier, Customer Service Department, 3 Killiney Road, #08-01 Winsland House I, Singapore 239519; phone: (+65) 6349 0222; fax: (+65) 6733 1510; e-mail: asiainfo@elsevier.com

© 2003 Elsevier Ltd. All rights reserved.

USA mailing notice: *Information Processing & Management* (ISSN 0306-4573) is published six issues per year in January, March, May, July, September and November by Elsevier Ltd. (P.O. Box 211, 1000 AE Amsterdam, The Netherlands). Annual subscription price in the USA US$ 1123.00 (valid in North, Central and South America), including air speed delivery. Application to mail at periodical postage rate is pending at Jamaica, NY 11431.

USA POSTMASTER: Send address changes to *Information Processing & Management*, Publications Expediting Inc., 200 Meacham Avenue, Elmont, NY 11003.

AIRFREIGHT AND MAILING in the USA by Publications Expediting Inc., 200 Meacham Avenue, Elmont, NY 11003.

LIST OF CONTENTS

Number 1

Zoran Stejić
Yasufumi Takama
Kaoru Hirota
1 Genetic algorithm-based relevance feedback for image retrieval using local similarity patterns

Oh-Woog Kwon
Jong-Hyeok Lee
25 Text categorization based on *k*-nearest neighbor approach for Web site classification

Akiko Aizawa
45 An information-theoretic perspective of tf–idf measures

Eric J. Stierna
Neil C. Rowe
67 Applying information-retrieval methods to software reuse: a case study

Jacques Savoy
75 Cross-language information retrieval: experiments based on CLEF 2000 corpora

Wann-Yun Shieh
Tien-Fu Chen
Jean Jyh-Jiun Shann
Chung-Ping Chung
117 Inverted file compression through document identifier reassignment

S.-C. Kao
H.-C. Chang
C.-H. Lin
133 Decision support for the academic library acquisition budget allocation via circulation database mining

Book reviews

Joan C. Durrance
149 *Evaluating Networked Information Services: Techniques, Policy, and Issues* by Charles R. McClure, John Carlo Bertot

Mike Thelwall
151 *Statistical Methods for the Information Professional: A Practical, Painless Approach to Understanding, Using, and Interpreting Statistics* by Liwen Vaughan

T. Saracevic
E. Voorhees
D. Harman
153 Letter to the editor

Karen Sparck Janes
156 Letter to the editor

Number 2

SPECIAL ISSUE
Modelling Vagueness and Subjectivity in Information Access

Fabio Crestani **Gabriella Pasi**	161	Handling vagueness, subjectivity, and imprecision in information access: an introduction to the special issue
Sandor Dominich	167	Connectionist interaction information retrieval
Sadaaki Miyamoto	195	Information clustering based on fuzzy multisets
Lynda Tamine **Claude Chrisment** **Mohand Boughanem**	215	Multiple query evaluation based on an enhanced genetic algorithm
Enrique Herrera-Viedma **Eduardo Peis**	233	Evaluating the informative quality of documents in SGML format from judgements by means of fuzzy linguistic techniques based on computing with words
D. Sánchez **J. Chamorro-Martínez** **M.A. Vila**	251	Modelling subjectivity in visual perception of orientation for image retrieval
Deok-Hwan Kim **Chin-Wan Chung**	267	Collection fusion using Bayesian estimation of a linear regression model in image databases on the Web
P. Bosc **L. Liétard** **O. Pivert**	287	Sugeno fuzzy integral as a basis for the interpretation of flexible queries involving monotonic aggregates
Danilo Montesi **Alberto Trombetta** **Peter A. Dearnley**	307	A similarity based relational algebra for Web and multimedia data
Paola Carrara **Giuseppe Fresta** **Anna Rampini**	323	Time-based retrieval of soft maps for environmental change detection
		Book reviews
Warren R. Greiff	329	Soft Computing in Information Retrieval: Techniques and Applications. F. Crestani and G. Pasi
P. Bosc	330	Uncertainty in Intelligent and Information Systems. B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh

Number 3

	333	Publisher's Note
Daniela Stan **Ishwar K. Sethi**	335	*e*ID: a system for exploration of image databases
Tadashi Nomoto **Yuji Matsumoto**	363	The diversity-based approach to open-domain text summarization
Ari Pirkola **Deniz Puolamäki** **Kalervo Järvelin**	391	Applying query structuring in cross-language retrieval
Howard Greisdorf	403	Relevance thresholds: a multi-stage predictive model of how users evaluate information
Mu-hsuan Huang	425	Pausal behavior of end-users in online searching
Pertti Vakkari **Mikko Pennanen** **Sami Serola**	445	Changes of search terms and tactics while writing a research proposal A longitudinal case study
Daniel P. Lorence **Amanda Spink**	465	Assessment of preferences for classification detail in medical information: is uniformity better?
David Nicholas **Paul Huntington** **Peter Williams**	479	Three years of digital consumer health information: a longitudinal study of the touch screen health kiosk
	503	Erratum to "Letters to the Editor"

Number 4

Bong-Hyun Cho **Changki Lee** **Gary Geunbae Lee**	505	Exploring term dependences in probabilistic information retrieval model
Hang Li **Kenji Yamanishi**	521	Topic analysis using a finite mixture model
P.L. Lin **W.H. Tan**	543	An efficient method for the retrieval of objects by topological relations in spatial database systems
Gary Marchionini **Xiangming Mu**	561	User studies informing E-table interfaces

Yves Rasolofo
David Hawking
Jacques Savoy
581 Result merging strategies for a current news metasearcher

Seda Ozmutlu
Amanda Spink
Huseyin C. Ozmutlu
611 Multimedia Web searching trends: 1997–2001

Dongseok Kim
Hanmin Jung
Gary Geunbae Lee
623 Unsupervised learning of mDTD extraction patterns for Web text mining

Bor-sheng Tsai
639 Information landscaping: information mapping, charting, querying and reporting techniques for total quality knowledge management

Book reviews

Hongmei Wang
Ya Zhang
665 Shaping Web Usability: Interaction Design in Context. By Albert N. Badre (Boston: Addison Wesley, 2002) [ISBN: 0201729938, pp. 304, Paperback, Price $39.99]

Jody M. Heckman
Edward J. Glantz
667 Web content management: a collaborative approach. By Nakano, Russell (2002). Web Content Management: A collaborative Approach, First Edition, Boston: Addison Wesley Professional [ISBN: 0-201-65782-1, 222 Pages, Paperback, Price: $39.99]

Number 5

Cristina López-Pujalte
Vicente P. Guerrero-Bote
Félix de Moya-Anegón
669 Genetic algorithms in relevance feedback: a second test and new contributions

Xia Lin
Howard D. White
Jan Buzydlowski
689 Real-time author co-citation mapping for online searching

Ryen W. White
Joemon M. Jose
Ian Ruthven
707 A task-oriented study on the influencing effects of query-biased summarisation in web searching

J. David Johnson
735 On contexts of information seeking

Quentin L. Burrell
761 Age-specific citation rates and the Egghe–Rao function

L. Egghe
C. Michel 771 Construction of weak and strong similarity measures for ordered sets of documents using fuzzy set techniques

Number 6

El-Sayed Atlam
M. Fuketa
K. Morita
Jun-ichi Aoe 809 Documents similarity measurement using field association terms

J. Morato
J. Llorens
G. Genova
J.A. Moreiro 825 Experiments in discourse analysis impact on information classification and retrieval algorithms

Peter Bailey
Nick Craswell
David Hawking 853 Engineering a multi-purpose test collection for Web retrieval experiments

Chen-Tung Chen
Wei-Shen Tai 873 An information push-delivery system design for personal information service on the Internet

Ziming Liu 889 Trends in transforming scholarly communication and their implications

Hong (Iris) Xie 899 Supporting ease-of-use and user control: desired features and structure of Web-based online IR systems

John E. Leide
Andrew Large
Jamshid Beheshti
Martin Brooks
Charles Cole 923 Visualization schemes for domain novices exploring a topic space: the navigation classification scheme

I List of Contents and Author Index